OIPE

RAW SEQUENCE LISTING   DATE: 02/06/2002
PATENT APPLICATION: US/09/813,271B   TIME: 12:05:36

Input Set : A:\20039SEQ.txt
Output Set: N:\CRF3\02062002\I813271B.raw

Does Not Comply
Corrected Diskette Needed

SEQUENCE LISTING
```
      4 (1) GENERAL INFORMATION:
      6       (i) APPLICANT:        𝒯 𝓂𝑜𝓋 𝓌𝓇
      7            _(A) Nico Cerletti
      9      (ii) TITLE OF INVENTION: New process for the production of
     10                              biologically active protein
     12     (iii) NUMBER OF SEQUENCES: 13
C--> 14      (iv) CORRESPONDENCE ADDRESS:
C--> 15            (A) ADDRESSEE: Novartis Patent Department
     16            (B) STREET: 564 Morris Avenue
     17            (C) CITY: Summit
     18            (D) STATE: New Jersey
     19            (E) COUNTRY: USA
     20            (F) ZIP: 07901
     22       (v) COMPUTER READABLE FORM:
     23            (A) MEDIUM TYPE: Floppy disk
     24            (B) COMPUTER: IBM PC compatible
     25            (C) OPERATING SYSTEM: PC-DOS/MS-DOS
     26            (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
     28      (vi) CURRENT APPLICATION DATA:
C--> 29            (A) APPLICATION NUMBER: US/09/813,271B
C--> 30            (B) FILING DATE: 20-Mar-2001
     32     (vii) PRIOR APPLICATION DATA:
     33            (A) APPLICATION NUMBER: PCT/EP95/02719
     34            (B) FILING DATE: 12-Jul-95
     35            (A) APPLICATION NUMBER: EPO 94810439.3
     36            (B) FILING DATE: 25-Jul-94
     38    (viii) ATTORNEY/AGENT INFORMATION:
     39            (A) NAME: Pfeiffer, Hesna J. .
     40            (B) REGISTRATION NUMBER: 22640
     41            (C) REFERENCE/DOCKET NUMBER: 4-20039C/C1C1/USN
     43      (ix) TELECOMMUNICATION INFORMATION:
     44            (A) TELEPHONE: (908) 522-6940
C--> 45            (B) TELEFAX: (908) 522-6955
     49 (2) INFORMATION FOR SEQ ID NO: 1:
     51       (i) SEQUENCE CHARACTERISTICS:
     52            (A) LENGTH: 339 base pairs
     53            (B) TYPE: nucleic acid
     54            (C) STRANDEDNESS: double
     55            (D) TOPOLOGY: linear
     57      (ii) MOLECULE TYPE: cDNA to mRNA
     59     (iii) HYPOTHETICAL: NO
     62     (vii) IMMEDIATE SOURCE:
```

RAW SEQUENCE LISTING   DATE: 02/06/2002
PATENT APPLICATION: US/09/813,271B   TIME: 12:05:36

Input Set : A:\20039SEQ.txt
Output Set: N:\CRF3\02062002\I813271B.raw

```
 63        (B) CLONE: E. coli LC137/pPLMu.hTGF-beta1 (DSM 5656)
 65    (ix) FEATURE:
 66        (A) NAME/KEY: CDS
 67        (B) LOCATION:1..336
 68        (D) OTHER INFORMATION:/product= "human TGF-beta1"
 71    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 73 GCC CTG GAC ACC AAC TAT TGC TTC AGC TCC ACG GAG AAG AAC TGC TGC          48
 74 Ala Leu Asp Thr Asn Tyr Cys Phe Ser Ser Thr Glu Lys Asn Cys Cys
 75   1             5                   10                  15
 77 GTG CGG CAG CTG TAC ATT GAC TTC CGC AAG GAC CTC GGC TGG AAG TGG          96
 78 Val Arg Gln Leu Tyr Ile Asp Phe Arg Lys Asp Leu Gly Trp Lys Trp
 79             20                  25                  30
 81 ATC CAC GAG CCC AAG GGC TAC CAT GCC AAC TTC TGC CTC GGG CCC TGC         144
 82 Ile His Glu Pro Lys Gly Tyr His Ala Asn Phe Cys Leu Gly Pro Cys
 83         35                  40                  45
 85 CCC TAC ATT TGG AGC CTG GAC ACG CAG TAC AGC AAG GTC CTG GCC CTG         192
 86 Pro Tyr Ile Trp Ser Leu Asp Thr Gln Tyr Ser Lys Val Leu Ala Leu
 87     50                  55                  60
 89 TAC AAC CAG CAT AAC CCG GGC GCC TCG GCG GCG CCG TGC TGC GTG CCG         240
 90 Tyr Asn Gln His Asn Pro Gly Ala Ser Ala Ala Pro Cys Cys Val Pro
 91   65                  70                  75                  80
 93 CAG GCG CTG GAG CCG CTG CCC ATC GTG TAC TAC GTG GGC CGC AAG CCC         288
 94 Gln Ala Leu Glu Pro Leu Pro Ile Val Tyr Tyr Val Gly Arg Lys Pro
 95             85                  90                  95
 97 AAG GTG GAG CAG CTG TCC AAC ATG ATC GTG CGC TCC TGC AAG TGC AGC         336
 98 Lys Val Glu Gln Leu Ser Asn Met Ile Val Arg Ser Cys Lys Cys Ser
 99             100                 105                 110
101 TGA                                                                     339
104 (2) INFORMATION FOR SEQ ID NO: 2:
106    (i) SEQUENCE CHARACTERISTICS:
107        (A) LENGTH: 112 amino acids
108        (B) TYPE: amino acid
109        (D) TOPOLOGY: linear
111    (ii) MOLECULE TYPE: protein
112    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
114 Ala Leu Asp Thr Asn Tyr Cys Phe Ser Ser Thr Glu Lys Asn Cys Cys
115   1             5                   10                  15
117 Val Arg Gln Leu Tyr Ile Asp Phe Arg Lys Asp Leu Gly Trp Lys Trp
118             20                  25                  30
120 Ile His Glu Pro Lys Gly Tyr His Ala Asn Phe Cys Leu Gly Pro Cys
121         35                  40                  45
123 Pro Tyr Ile Trp Ser Leu Asp Thr Gln Tyr Ser Lys Val Leu Ala Leu
124     50                  55                  60
126 Tyr Asn Gln His Asn Pro Gly Ala Ser Ala Ala Pro Cys Cys Val Pro
127   65                  70                  75                  80
129 Gln Ala Leu Glu Pro Leu Pro Ile Val Tyr Tyr Val Gly Arg Lys Pro
130             85                  90                  95
132 Lys Val Glu Gln Leu Ser Asn Met Ile Val Arg Ser Cys Lys Cys Ser
133             100                 105                 110
```

```
                RAW SEQUENCE LISTING            DATE: 02/06/2002
                PATENT APPLICATION: US/09/813,271B      TIME: 12:05:36

                Input Set : A:\20039SEQ.txt
                Output Set: N:\CRF3\02062002\I813271B.raw


136 (2) INFORMATION FOR SEQ ID NO: 3:
138      (i) SEQUENCE CHARACTERISTICS:
139            (A) LENGTH: 339 base pairs
140            (B) TYPE: nucleic acid
141            (C) STRANDEDNESS: double
142            (D) TOPOLOGY: linear
144     (ii) MOLECULE TYPE: cDNA to mRNA
147    (vii) IMMEDIATE SOURCE:
148            (B) CLONE: E. coli LC137/pPLMu.hTGF-beta2 (DSM5657)
150     (ix) FEATURE:
151            (A) NAME/KEY: CDS
152            (B) LOCATION:1..336
153            (D) OTHER INFORMATION:/product= "human TGF-beta2"
156     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
158 GCT TTG GAT GCG GCC TAT TGC TTT AGA AAT GTG CAG GAT AAT TGC TGC      48
159 Ala Leu Asp Ala Ala Tyr Cys Phe Arg Asn Val Gln Asp Asn Cys Cys
W--> 160        115                 120                 125
162 CTA CGT CCA CTT TAC ATT GAT TTC AAG AGG GAT CTA GGG TGG AAA TGG      96
163 Leu Arg Pro Leu Tyr Ile Asp Phe Lys Arg Asp Leu Gly Trp Lys Trp
W--> 164        130                 135                 140
166 ATA CAC GAA CCC AAA GGG TAC AAT GCC AAC TTC TGT GCT GGA GCA TGC     144
167 Ile His Glu Pro Lys Gly Tyr Asn Ala Asn Phe Cys Ala Gly Ala Cys
W--> 168 145                150                 155                 160
170 CCG TAT TTA TGG AGT TCA GAC ACT CAG CAC AGC AGG GTC CTG AGC TTA     192
171 Pro Tyr Leu Trp Ser Ser Asp Thr Gln His Ser Arg Val Leu Ser Leu
W--> 172                165                 170                 175
174 TAT AAT ACC ATA AAT CCA GAA GCA TCT GCT TCT CCT TGC TGC GTG TCC     240
175 Tyr Asn Thr Ile Asn Pro Glu Ala Ser Ala Ser Pro Cys Cys Val Ser
W--> 176                180                 185                 190
178 CAA GAT TTA GAA CCT CTA ACC ATT CTC TAC TAC ATT GGC AAA ACA CCC     288
179 Gln Asp Leu Glu Pro Leu Thr Ile Leu Tyr Tyr Ile Gly Lys Thr Pro
W--> 180        195                 200                 205
182 AAG ATT GAA CAG CTT TCT AAT ATG ATT GTA AAG TCT TGC AAA TGC AGC     336
183 Lys Ile Glu Gln Leu Ser Asn Met Ile Val Lys Ser Cys Lys Cys Ser
W--> 184 210                215                 220
186 TAA                                                                 339
189 (2) INFORMATION FOR SEQ ID NO: 4:
191      (i) SEQUENCE CHARACTERISTICS:
192            (A) LENGTH: 112 amino acids
193            (B) TYPE: amino acid
194            (D) TOPOLOGY: linear
196     (ii) MOLECULE TYPE: protein
197     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
199 Ala Leu Asp Ala Ala Tyr Cys Phe Arg Asn Val Gln Asp Asn Cys Cys
200   1                 5                  10                  15
202 Leu Arg Pro Leu Tyr Ile Asp Phe Lys Arg Asp Leu Gly Trp Lys Trp
203              20                 25                  30
205 Ile His Glu Pro Lys Gly Tyr Asn Ala Asn Phe Cys Ala Gly Ala Cys
206          35                 40                  45
```

```
              RAW SEQUENCE LISTING              DATE: 02/06/2002
              PATENT APPLICATION:  US/09/813,271B     TIME: 12:05:36

              Input Set : A:\20039SEQ.txt
              Output Set: N:\CRF3\02062002\I813271B.raw


208 Pro Tyr Leu Trp Ser Ser Asp Thr Gln His Ser Arg Val Leu Ser Leu
209         50              55              60
211 Tyr Asn Thr Ile Asn Pro Glu Ala Ser Ala Ser Pro Cys Cys Val Ser
212 65              70              75              80
214 Gln Asp Leu Glu Pro Leu Thr Ile Leu Tyr Tyr Ile Gly Lys Thr Pro
215             85              90              95
217 Lys Ile Glu Gln Leu Ser Asn Met Ile Val Lys Ser Cys Lys Cys Ser
218         100             105             110
221 (2) INFORMATION FOR SEQ ID NO: 5:
223      (i) SEQUENCE CHARACTERISTICS:
224          (A) LENGTH: 339 base pairs
225          (B) TYPE: nucleic acid
226          (C) STRANDEDNESS: double
227          (D) TOPOLOGY: linear
229     (ii) MOLECULE TYPE: cDNA to mRNA
232    (vii) IMMEDIATE SOURCE:
233          (B) CLONE: E. coli LC137/pPLMu.hTGF-beta3 (DSM 5658)
235     (ix) FEATURE:
236          (A) NAME/KEY: CDS
237          (B) LOCATION:1..336
238          (D) OTHER INFORMATION:/product= "human TGF-beta3"
241     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
243 GCT TTG GAC ACC AAT TAC TGC TTC CGC AAC TTG GAG GAG AAC TGC TGT        48
244 Ala Leu Asp Thr Asn Tyr Cys Phe Arg Asn Leu Glu Glu Asn Cys Cys
W--> 245         115             120             125
247 GTG CGC CCC CTC TAC ATT GAC TTC CGA CAG GAT CTG GGC TGG AAG TGG        96
248 Val Arg Pro Leu Tyr Ile Asp Phe Arg Gln Asp Leu Gly Trp Lys Trp
W--> 249     130             135             140
251 GTC CAT GAA CCT AAG GGC TAC TAT GCC AAC TTC TGC TCA GGC CCT TGC       144
252 Val His Glu Pro Lys Gly Tyr Tyr Ala Asn Phe Cys Ser Gly Pro Cys
W--> 253 145             150             155             160
255 CCA TAC CTC CGC AGT GCA GAC ACA ACC CAC AGC ACG GTG CTG GGA CTG       192
256 Pro Tyr Leu Arg Ser Ala Asp Thr Thr His Ser Thr Val Leu Gly Leu
W--> 257             165             170             175
259 TAC AAC ACT CTG AAC CCT GAA GCA TCT GCC TCG CCT TGC TGC GTG CCC       240
260 Tyr Asn Thr Leu Asn Pro Glu Ala Ser Ala Ser Pro Cys Cys Val Pro
W--> 261             180             185             190
263 CAG GAC CTG GAG CCC CTG ACC ATC CTG TAC TAT GTT GGG AGG ACC CCC       288
264 Gln Asp Leu Glu Pro Leu Thr Ile Leu Tyr Tyr Val Gly Arg Thr Pro
W--> 265         195             200             205
267 AAA GTG GAG CAG CTC TCC AAC ATG GTG GTG AAG TCT TGT AAA TGT AGC       336
268 Lys Val Glu Gln Leu Ser Asn Met Val Val Lys Ser Cys Lys Cys Ser
W--> 269 210             215             220
271 TGA                                                                   339
274 (2) INFORMATION FOR SEQ ID NO: 6:
276      (i) SEQUENCE CHARACTERISTICS:
277          (A) LENGTH: 112 amino acids
278          (B) TYPE: amino acid
279          (D) TOPOLOGY: linear
```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/813,271B
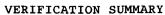
DATE: 02/06/2002
TIME: 12:05:36

Input Set : A:\20039SEQ.txt
Output Set: N:\CRF3\02062002\I813271B.raw

```
281      (ii) MOLECULE TYPE: protein
282      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
284 Ala Leu Asp Thr Asn Tyr Cys Phe Arg Asn Leu Glu Glu Asn Cys Cys
285  1               5                  10                  15
287 Val Arg Pro Leu Tyr Ile Asp Phe Arg Gln Asp Leu Gly Trp Lys Trp
288             20                  25                  30
290 Val His Glu Pro Lys Gly Tyr Tyr Ala Asn Phe Cys Ser Gly Pro Cys
291          35                  40                  45
293 Pro Tyr Leu Arg Ser Ala Asp Thr Thr His Ser Thr Val Leu Gly Leu
294       50                  55                  60
296 Tyr Asn Thr Leu Asn Pro Glu Ala Ser Ala Ser Pro Cys Cys Val Pro
297  65                  70                  75                  80
299 Gln Asp Leu Glu Pro Leu Thr Ile Leu Tyr Tyr Val Gly Arg Thr Pro
300                  85                  90                  95
302 Lys Val Glu Gln Leu Ser Asn Met Val Val Lys Ser Cys Lys Cys Ser
303                 100                 105                 110
306 (2) INFORMATION FOR SEQ ID NO: 7:
308      (i) SEQUENCE CHARACTERISTICS:
309          (A) LENGTH: 336 base pairs
310          (B) TYPE: nucleic acid
311          (C) STRANDEDNESS: double
312          (D) TOPOLOGY: linear
314      (ii) MOLECULE TYPE: other nucleic acid
315          (A) DESCRIPTION: /desc = "recombinant hybrid DNA of
316                  TGF-beta1 and TGF-beta3 DNA"
319      (vii) IMMEDIATE SOURCE:
320          (B) CLONE: E. coli LC137/pPLMu.TGF-beta1(44/45)beta3
322      (ix) FEATURE:
323          (A) NAME/KEY: mat_peptide
324          (B) LOCATION:1..132
325          (D) OTHER INFORMATION:/product= "N-terminal 44 amino
326 acids of human TGF-beta1"
328      (ix) FEATURE:
329          (A) NAME/KEY: mat_peptide
330          (B) LOCATION:133..336
331          (D) OTHER INFORMATION:/product= "C-terminal 68 amino
332 acids of human TGF-beta3"
334      (ix) FEATURE:
335          (A) NAME/KEY: CDS
336          (B) LOCATION:1..336
337          (D) OTHER INFORMATION:/product= "hybrid TGF-beta named
338 TGF-beta1-3"
341      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:
343 GCC CTG GAC ACC AAC TAT TGC TTC AGC TCC ACG GAG AAG AAC TGC TGC      48
344 Ala Leu Asp Thr Asn Tyr Cys Phe Ser Ser Thr Glu Lys Asn Cys Cys
345  1               5                  10                  15
347 GTG CGG CAG CTG TAC ATT GAC TTC CGC AAG GAC CTC GGC TGG AAG TGG      96
348 Val Arg Gln Leu Tyr Ile Asp Phe Arg Lys Asp Leu Gly Trp Lys Trp
349             20                  25                  30
```

W--> 316 (marginal note pointing to line 316)

VERIFICATION SUMMARY   DATE: 02/06/2002
PATENT APPLICATION: US/09/813,271B   TIME: 12:05:37

Input Set : **A:\20039SEQ.txt**
Output Set: **N:\CRF3\02062002\I813271B.raw**

```
L:14  M:220 C: Keyword misspelled or invalid format, [(iv) CORRESPONDENCE ADDRESS:]
L:15  M:220 C: Keyword misspelled or invalid format, [(A) ADDRESSEE:]
L:29  M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:30  M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:45  M:220 C: Keyword misspelled or invalid format, [(B) TELEFAX:]
L:160 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:164 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:168 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:172 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:176 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:180 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:184 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:245 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:5
L:249 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:5
L:253 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:5
L:257 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:5
L:261 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:5
L:265 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:5
L:269 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:5
```